Urinary microbiome of kidney transplant patients reveals dysbiosis with potential for antibiotic resistance

Asha Rani, PhD[1$], Ravi Ranjan PhD[1$], Halvor S. McGee, PhD[1], Kalista E. Andropolis, MS[1], Dipti V. Panchal, MS[1], Zahraa Hajjiri, MD[1], Daniel C. Brennan, MD[2], Patricia W. Finn, MD[1], David L. Perkins, MD, PhD[1,3,*]

[1]Department of Medicine, University of Illinois, Chicago, IL 60612 USA
[2]Division of Renal Diseases, Washington University School of Medicine, St. Louis, MO 63110 USA
[3]Department of Surgery, University of Illinois, Chicago, IL 60612 USA

[*]**Corresponding author:**
David Perkins, MD, PhD
University of Illinois at Chicago
Department of Medicine, MC 787
840 S Wood Street, Suite 1020N CSB
Chicago IL 60612 USA
Email: perkinsd@uic.edu , Phone: 312-413-3382, Fax: 312-355-0499

[$]These authors contributed equally and considered as co-first authors.

RUNNING HEAD
Urinary microbiome of kidney transplant patients

SUPPORTING FIGURE
Figure S1 to S7

SUPPORTING TABLE
Table S1 to S6

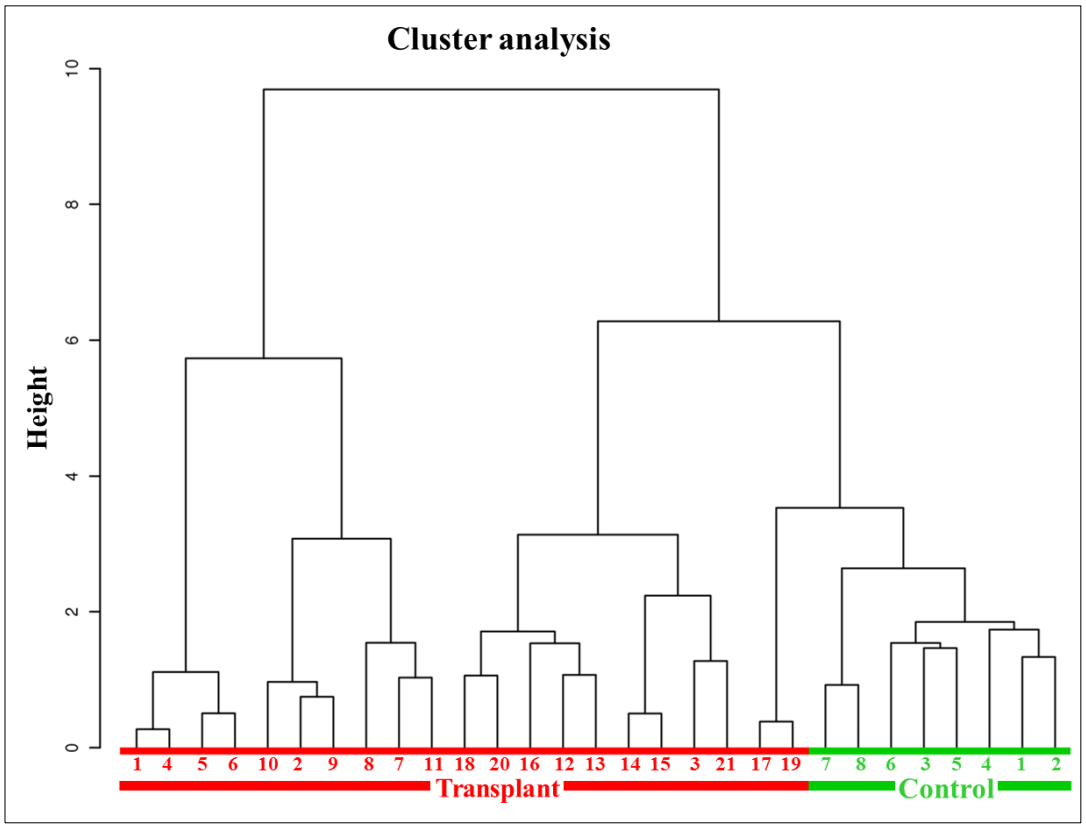

Figure S1. Dendrogram of the urine microbiome. The dendrogram is based on hierarchical clustering of transplant and control groups using the Ward method with the Pearson's correlation distance metric.

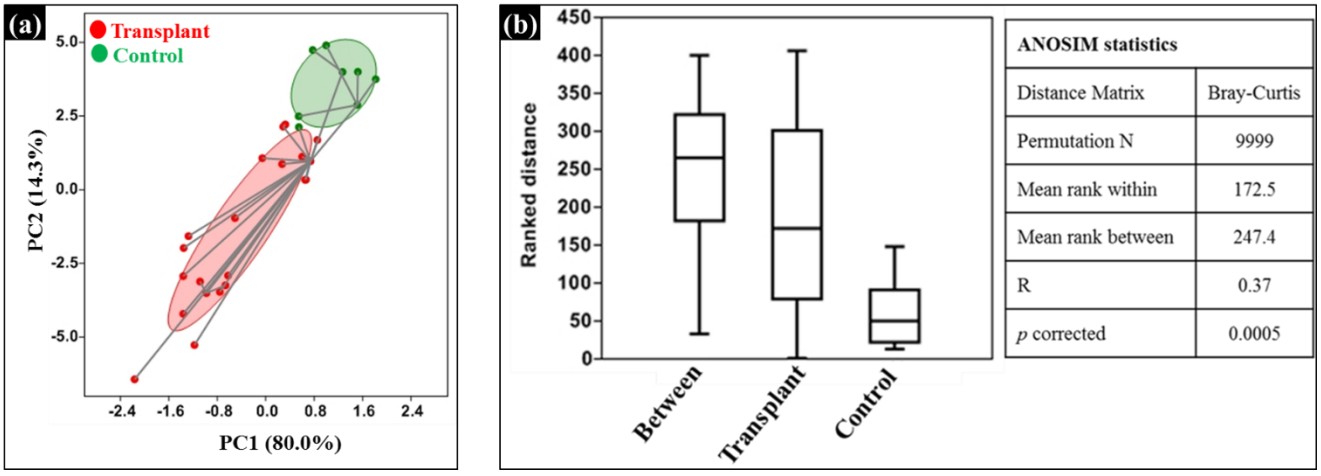

Figure S2. Multivariate analysis of microbial diversity. (a) Partial least squares discriminant analysis (PLS-DA) of the bacterial species in transplant and control group. The percentage of total variance explained by PC1 and PC2 are noted in both the axis labels. (b) ANOSIM (Analysis of Similarity) test showed that the within-group variability was low compared to between-group variability.

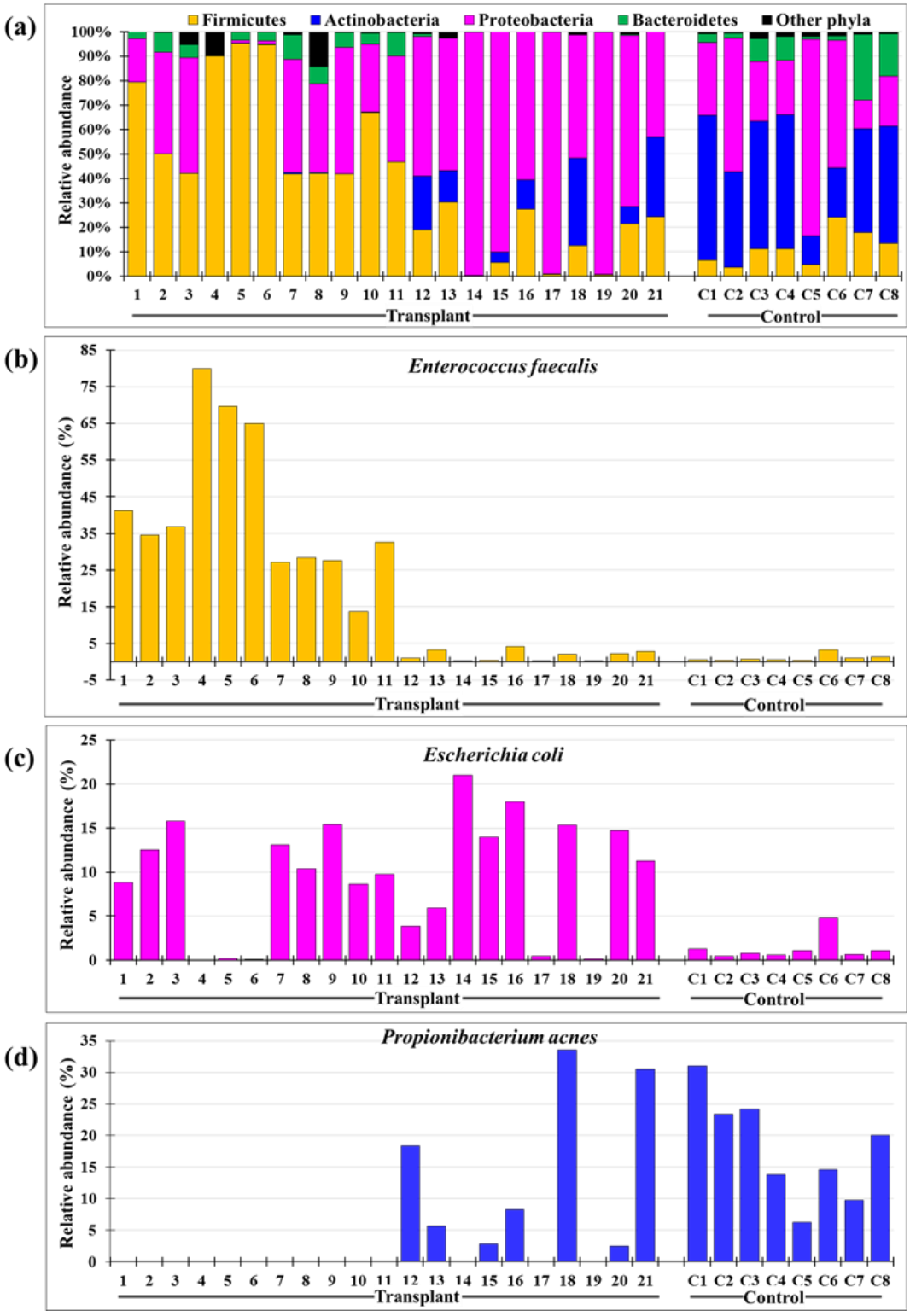

Figure S3. Microbial diversity profile of transplant and control group. (a) Stacked bar chart representation of relative abundance of four dominant phyla. Relative abundance of *E. faecalis* (b), *E. coli* (c), and *P. acnes* (d) for individual samples in transplant and control group. The color of the bar graph (species) represent the phylum which they belong (stacked bar graph).

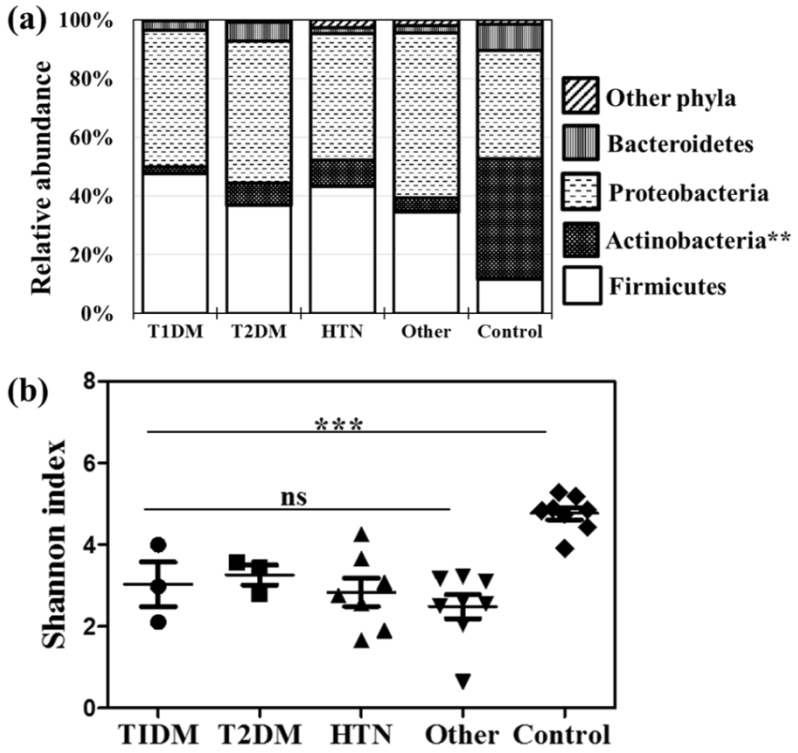

Figure S4. Bacterial phyla and microbial diversity in different diagnostic groups. (a) Relative abundances of the 4 dominant bacterial phyla detected in different transplant diagnoses compared to control. The other phyla category contains phyla with < 2% abundance. The significance of the difference was computed using ANOVA with Tukey's multiple comparison test. Asterisks indicate significant differences at phylum level (**p<0.001). **(b)** Shannon diversity index in 4 clinical diagnoses compared to control, and is significantly different in control group compared to 4 transplant diagnosis categories. No significant differences were observed among the 4 transplant diagnosis categories. "Other" includes the combined category for FSGS, PKD, RN, Lupus and vasculitis (Table 1). The significance of the difference was computed using ANOVA with Tukey's multiple comparison test. Asterisks indicate significant differences at species level (***p<0.0001), see Supplementary Table 6.

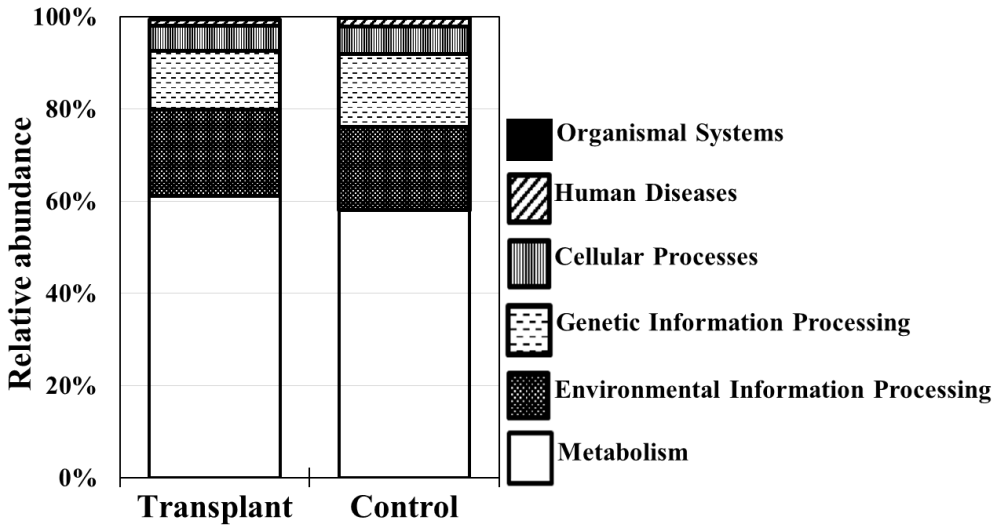

Figure S5. No change in the level 1 KEGG functional pathway categories between the transplant and control urine. Shown are relative abundance of KEGG assignments for the level 1 KEGG categories identified in transplant and control group. Both groups are highly enriched for metabolism and environmental information processing categories. No significant differences between the groups were observed at Level 1.

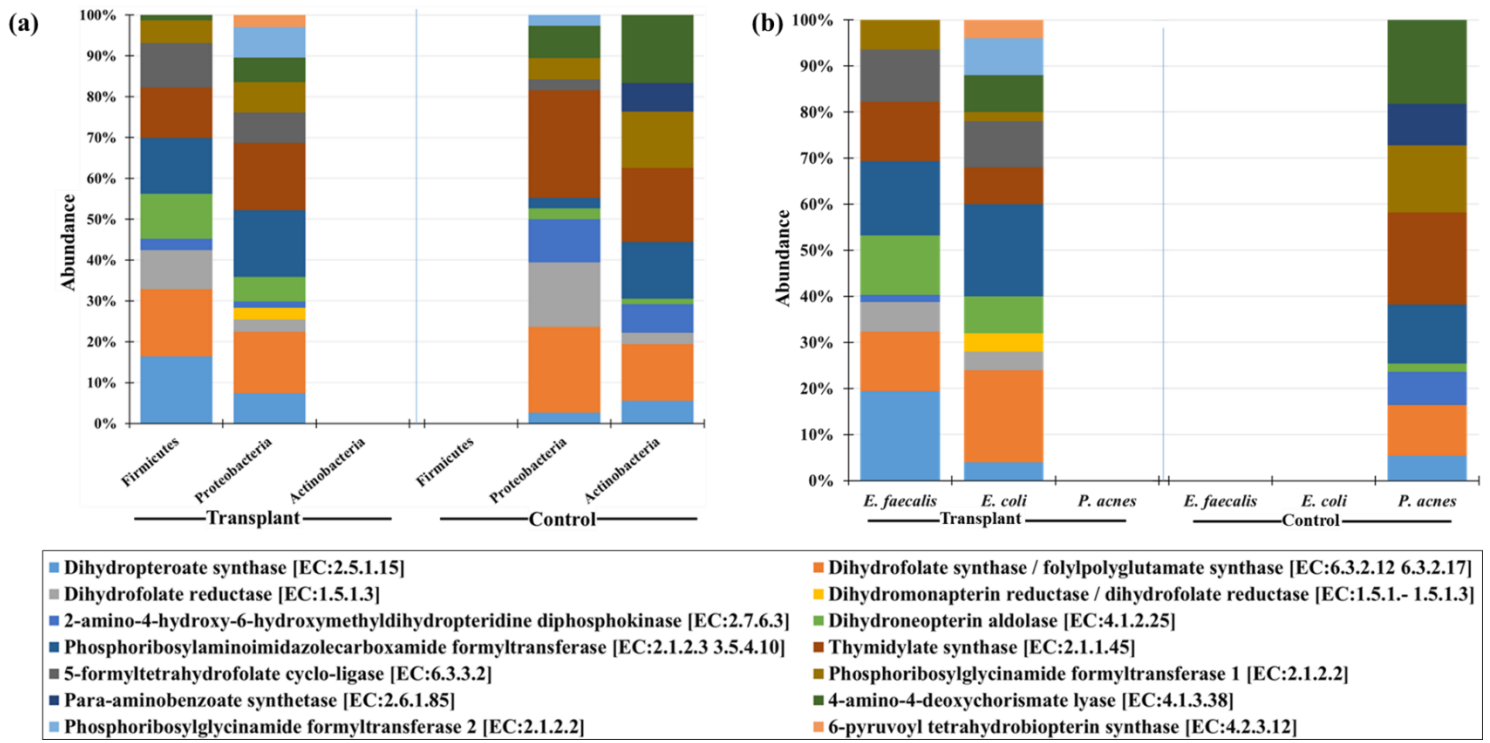

Figure S6. Abundance of folate biosynthesis pathway genes identified in transplant and control group. (a) Relative abundance of folate pathway genes compared among; (a) major phyla – Firmicutes, Proteobacteria, and Actinobacteria, **(b)** Bacterial species – *E. faecalis*, *E. coli*, and *P. acnes* representing the respective phyla.

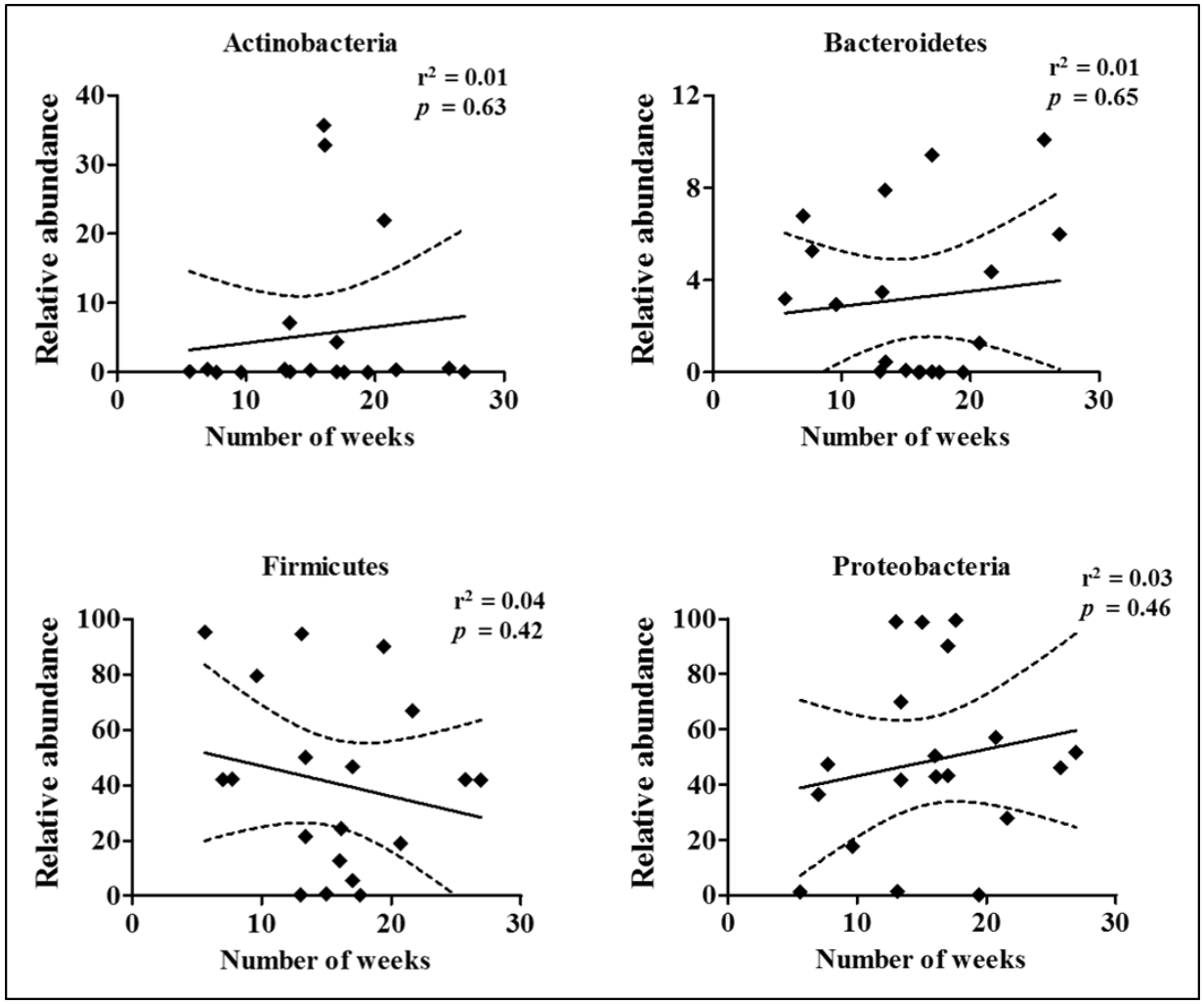

Figure S7. Correlation analysis. The correlation analysis using linear regression of sample collection time points with the relative abundance of major bacterial phyla identified in the urine samples. No significant correlation was observed.

Table S1. Statistics of metagenomics data.

Sample name	Number of reads	Reads used for contig assembly	No. of contigs
1	7849916	7814972	47
2	21803762	5819208	2471
3	10327248	10325043	25
4	12785361	12039222	14
5	7433295	7419132	374
6	8174663	8150995	312
7	34818382	3134361	7487
8	27420614	4565511	4935
9	31903768	6576452	2417
10	28570804	8361721	6376
11	35156390	5717893	5503
12	642591	503045	2661
13	392172	310036	3306
14	281816	171627	4799
15	399222	298956	1992
16	580114	464905	2164
17	689472	519830	8874
18	881827	656455	3234
19	446857	324868	8414
20	563269	485818	1604
21	295740	225778	1611
C1	7545864	5145175	10732
C2	5384830	4007780	14516
C3	9235514	5503626	15563
C4	9790696	5349040	16247
C5	8187139	5957934	13343
C6	8043242	4598491	5387
C7	4690991	2233658	3041
C8	3756888	1830113	2898
Group	**Number of reads**	**Reads used for analysis**	**Total number of contigs analyzed**
Transplant	11019870 ± 2852048	3995000 ± 860800	68620
Control	7079396 ± 779065	4328000 ± 543600	81727

Total number of reads include raw sequence reads for each sample. Reads used for analysis includes sequences retained after quality filter and human sequence removal. Number of reads for each group are shown as mean ± SEM.

Table S2. **Abundance of bacterial genera.**

Transplant	
Genus	**% Abundance**
Enterococcus*	**28.62**
Escherichia*	**12.67**
Ralstonia	**9.56**
*Shigella**	7.43
Propionibacterium*	**4.85**
Proteus	4.45
Streptococcus	**3.41**
*Lactobacillus**	2.85
Bacteroides	**2.40**
Neisseria	**2.22**
Salmonella	1.64
Burkholderia	**1.30**
*Staphylococcus**	1.17
*Citrobacter**	1.12

Control	
Genus	**% Abundance**
Propionibacterium*	**18.48**
*Corynebacterium**	10.55
Ralstonia	**7.58**
Neisseria	**7.27**
Streptococcus	**3.23**
Prevotella	2.65
Bacteroides	**2.55**
*Mobiluncus**	2.44
Porphyromonas	2.20
Enterococcus*	**2.18**
Pseudomonas	2.11
Burkholderia	**1.65**
Enhydrobacter	1.51
Escherichia*	**1.39**

Bacterial genera identified in transplant and control group (selected on the basis of >1% abundance in each group). Genus identified in both the groups are shown in bold letters, whereas non-bold are the genera specific to each group. The significance of the difference among bacterial genera was computed using Wilcoxon Rank test (*p corrected < 0.05).

Table S3. Richness and diversity measures of urine microbiome.

Diversity estimators	Transplant		Control	
	Mean	Min–Max	Mean	Min–Max
Shannon index***	2.8±0.2	0.6-4.3	4.8±0.2	3.9-5.3
Inverse Simpson index**	0.8±0.0	0.3-0.9	0.9±0.0	0.8-0.9
Evenness	0.6±0.0	0.3-0.9	0.7±0.0	0.6-0.8

Richness and diversity measures of urine microbiome including the Shannon index, Inverse Simpson index, and evenness are shown. The index values are shown as group mean ± SD. The significance of the difference was computed using Mann Whitney *U* test (**$p < 0.001$, ***$p < 0.0001$).

Table S4. Abundance of bacterial species in each diagnostic category.

TIDM	
Species	**% Abundance**
Enterococcus faecalis	23.64
Escherichia coli	12.97
Enterococcus sp.	4.10
Enterococcus sp. 7L76	3.89
Bacteroides vulgatus	2.86
Salmonella enterica	2.83
Shigella dysenteriae	2.29
Escherichia sp. 4_1_40B	2.26
Ralstonia pickettii	2.18
Shigella sonnei	1.79
Shigella flexneri	1.65
Streptococcus pneumoniae	1.64
Enterococcus faecium	1.59
Citrobacter koseri	1.56
Shigella boydii	1.54
Citrobacter sp. 30_2	1.36
Shigella sp. D9	1.15
Shigella sp.	1.15
Streptococcus sanguinis	1.13
Klebsiella pneumoniae	1.02

T2DM	
Species	**% Abundance**
Enterococcus faecalis	20.91
Escherichia coli	9.80
Propionibacterium acnes	6.14
Ralstonia pickettii	4.60
Bacteroides vulgatus	4.39
Escherichia sp. 4_1_40B	3.45
Enterococcus sp.	3.21
Enterococcus sp. 7L76	2.50
Ralstonia solanacearum	1.88
Shigella flexneri	1.86
Shigella boydii	1.85
Shigella dysenteriae	1.77
Shigella sp. D9	1.77
Shigella sonnei	1.76
Shigella sp.	1.76
Staphylococcus epidermidis	1.32
Ralstonia sp. 5_7_47FAA	1.28
Ralstonia sp.	1.28
Enterococcus faecium	1.09
Escherichia sp. 1_1_43	1.00

HTN	
Species	**% Abundance**
Enterococcus faecalis	24.63
Escherichia coli	7.21
Propionibacterium acnes	6.80
Ralstonia pickettii	5.59
Ralstonia solanacearum	4.47
Enterococcus faecium	2.94
Bacteroides vulgatus	1.79
Salmonella enterica	1.62
Enterococcus sp.	1.57
Ralstonia sp.	1.35
Mycoplasma hominis	1.31
Pseudomonas aeruginosa	1.31
Ralstonia sp. 5_7_47FAA	1.30
Proteus mirabilis	1.28
Enterococcus sp. 7L76	1.17
Streptococcus agalactiae	0.95
Neisseria meningitidis	0.94
Streptococcus pyogenes	0.93
Streptococcus pneumoniae	0.74
Escherichia sp. 4_1_40B	0.71

Other	
Species	**% Abundance**
Enterococcus faecalis	20.83
Escherichia coli	10.08
Proteus mirabilis	9.69
Propionibacterium acnes	4.16
Ralstonia pickettii	3.80
Ralstonia solanacearum	3.62
Lactobacillus crispatus	3.07
Bacteroides vulgatus	1.92
Shigella sonnei	1.79
Shigella sp. D9	1.61
Shigella dysenteriae	1.60
Escherichia sp. 4_1_40B	1.59
Shigella sp.	1.59
Enterococcus faecium	1.47
Salmonella enterica	1.36
Shigella flexneri	1.35
Bacterium RRLBTPL IV-3	1.25
Lactobacillus acidophilus	1.19
Ralstonia sp.	1.14
Shigella boydii	1.13

Shown are bacterial species identified in 4 different clinical diagnosis categories including T1DM (Type 1 Diabetes Mellitus); T2DM (Type 2 Diabetes Mellitus); HTN (Hypertension) and "Other" includes the combined category for FSGS, Nephrocalcinosis, PKD, RN, Lupus and vasculitis (Table 1). The first 20 species are selected on the basis of >1% abundance in the clinical diagnosis category "Other". Bacterial species identified in at least 3 diagnostic category are shaded. Statistical analysis was performed using ANOVA with Tukey's multiple comparison post-test. No significant differences were observed.

Table S5. Statistics of the abundance of bacterial species in each diagnostic category compared with controls.

Table analyzed	**Comparison at the species level in 4 clinical diagnosis categories with control**				
One-way analysis of variance					
P value	< 0.0001				
P value summary	***				
Are means signif. different (P < 0.05)	Yes				
Number of groups	5				
F	10.74				
R squared	0.64				
ANOVA Table	SS	df	MS		
Treatment (between columns)	24.19	4	6.05		
Residual (within columns)	13.52	24	0.56		
Total	37.71	28			
Tukey's Multiple Comparison Test	**Mean Diff.**	**q**	**Significant P < 0.05**	**Summary**	**95% CI of diff**
TIDM vs T2DM	-0.23	0.53	No	ns	-2.037 to 1.573
TIDM vs HTN	0.19	0.51	No	ns	-1.338 to 1.713
TIDM vs Other	0.55	1.52	No	ns	-0.9516 to 2.042
T2DM vs HTN	0.42	1.15	No	ns	-1.106 to 1.945
T2DM vs Other	0.78	2.16	No	ns	-0.7199 to 2.273
HTN vs Other	0.36	1.30	No	ns	-0.7867 to 1.502
TIDM vs Control	-1.74	4.84	Yes	*	-3.234 to -0.2407
T2DM vs Control	-1.51	4.19	Yes	*	-3.002 to -0.009023
HTN vs Control	-1.93	7.01	Yes	***	-3.069 to -0.7809
Other vs Control	-2.28	8.60	Yes	***	-3.388 to -1.177

Shannon index is significantly different in control group compared to 4 transplant diagnosis categories. No significant differences were observed among the 4 transplant diagnosis categories. Different clinical diagnosis categories are T1DM (Type 1 Diabetes Mellitus); T2DM (Type 2 Diabetes Mellitus); HTN (Hypertension) and "Other" includes the combined category for FSGS, Nephrocalcinosis, PKD, RN, Lupus and vasculitis (Table 1). The significance of the difference was computed using ANOVA with Tukey's multiple comparison test. Asterisks indicate significant differences at species level (*p<0.05, ***p<0.0001).

Table S6. Statistics of the abundance of KEGG Level 1 categories.

Level 1 KEGG categories	Transplant	Control
Metabolism	61.1	58.0
Environmental information processing	18.7	18.2
Genetic information processing	12.7	15.7
Cellular processes	5.4	5.9
Human diseases	1.3	1.9
Organismal systems	0.7	0.3

Shown are relative abundance of KEGG assignments identified in transplant and control group. Using Wilcoxon Rank test (*p* corrected <0.05) no significant differences between the metabolic categories were observed at Level 1.